QA
76
M94
SIRef

M U D P I E no. 14

Museum and University Data, Program and Information Exchange

COMPUTERS IN PALEONTOLOGY

Part B of the Proceedings of the North American Paleontological Convention, held in Chicago in 1969, has been published under the auspices of Allen Press, and can be purchased from them for $3.50 (Allen Press, Lawrence, Kansas, 66045). Six papers are included in this symposium. The titles are listed below to give a clue as to what was covered:

J. F. Mello. Paleontologic data storage and retrieval. pp. 57-70.

D. M. Raup. Modeling and simulation of morphology by computer. pp. 71-83.

R. L. Kaesler. Numerical taxonomy in paleontology: classification, ordination and reconstruction of phylogenies. pp. 84-100.

M. A. Buzas. On the quantification of biofacies. pp. 101-16.

W. T. Fox. Analysis and simulation of paleoecologic communities through time. pp. 117-134.

E. C. Olson. Current and projected impacts of computers upon concepts and research in paleontology. pp. 135-53.

AVAILABLE PROGRAMS

PITFALL I--A general purpose data processing program for environmental data, written by N. C. Hull, J. J. Beauchamp, and D. E. Reichle. Designed for the processing and analysis of complicated data sets which need storage, recall, summarization and basic statistical analysis. Developed to handle ecological data obtained in environmental monitoring. Handles any number of parameters, although will process only two simultaneously. Especially useful for problems of environmental quality and biological productivity. Includes three integrated programs; one to transfer raw data from cards to mag tape, one to combine, recall, and print raw data with statistics, and one to permit statistical evaluation of summarized data sets using analysis of variance. Available from Clearinghouse for Federal Scientific and Technical Information, Springfield, Va. 22151, for $3.00. Refer. no.: CRNL-IBP-70-2.

[UNNAMED]--A FORTRAN IV program, written by G. L. Baskerville, for judging homogeneity of variance around regression for data untransformed or transformed by logarithms. Available as above from Clearing House, for $3.00. Refer. no.: ORNL-IBP-70-1.

CHEESE, ANYONE?

I had thought the ultimate in "buy-four-get-one-free" had been reached when I read about the Exotic Cheese Club, with its gruyere of the month, but I've just been invited to join the "Library of Computer and Information Sciences," which offers to help you keep "pace with the giant strides in computer science." They claim to screen hundreds of manuscripts each month in choosing only the most informative and important to offer to over 22,000 members. They do list some interesting stuff, of course, and in case any MUDPIE reader might like this easy way to "keep up-to-date and move ahead in their careers" (yes, that is what it said), you can contact them at the Library of Computer and Information Sciences, Riverside, N. J., 08075.--JAP.

RESOURCE SHARING COMPUTER NETWORKS

The Advanced Research Projects Agency of the Department of Defense, Washington, D. C., 20301, has put together a collection of papers presented at the Spring Joint Computer Conference, Atlantic City, N. J., May, 1970, under the above title. All papers are concerned with aspects of the ARPA Computer Network, which is the most ambitious attempt to set up a nationwide computer interchange to date. The problems are multiple and staggering, and it is a good thing we've got this group working on them. At the present time, participation in the network is restricted to DOD grantees engaged in computer research. As presently envisioned, the network will be expensive for anyone who wishes to join, and it will be out of the reach of museums for a long time--but it doesn't hurt to keep up with what they are doing. The papers included in this special leaflet were all published in the report of the proceedings of the Spring Joint Computer Conference, 1970, and the pagination included below is from that publication.-JAP.

Roberts, L. G. & B. D. Wessler. Computer network development to achieve resource sharing. [pp. 543-549.]

Heart, F. E., et al. The interface message processor for the ARPA computer network. [pp. 551-567.]

Kleinrock, L. Analytic and simulation methods in computer network design. [pp. 569-579.]

Frank, H., I. T. Frisch & W. Chou. Topological considerations in the design of the ARPA computer network. [pp. 581-7.]

Carr, C. S.ET AL. HOST-HOST communication protocol in the ARPA network. [pp. 589-597.]

INFORMATION SYSTEMS SURVEY
"TAXIR"
[Taxonomic Information Retrieval System]

TAXIR is a specialized information processing program with a high level language interface for defining simple queries or for requesting preparation of a catalog based upon the complete content of a data base.

General applicability is limited by two key factors. It is designed to handle static data bases rather than those which grow or change frequently because the data matrix which is generated by TAXIR and is an integral part of the system requires the availability of all data records at the time it is created. In addition, the data matrix must be core resident during query processing, hence limiting the size of the data base which can be handled.

The system was built at the University of Colorado under Dr. David Rogers. It is written in a high level language and is currently operational on several types of computers.

FILE ORGANIZATION -- The TAXIR data file is made up of a collection of records each of which is identical in format to the other records in the collection. Thus each record is made up of a fixed set of fields. No hierarchical data organization is permitted. Each field may assume one of a predefined set of values. This file is stored sequentially.

This data base is augmented by a highly compacted matrix which represents the content of each record and which is used to identify relevant records as defined by the selection criteria. This matrix is referred to as a construct and must be completely core resident when using TAXIR as a query system.

The degree of data independence built into the system is not known at this time.

FILE CREATION -- A complete data base is created in one run based upon input made up of a data string for each record with one entry per field. A construct is developed and stored for the data base. No automatic error analysis on content is provided by the system.

FILE MAINTENANCE -- Not applicable since the complete data base must be present at file creation.

RETRIEVAL -- Retrieval is controlled by the user through a high level language composed of a Boolean selection statement made up of elements of the form: data item = data value. No parentheticals may be used.

Search is performed on the construct and hence is rapid.

Tabulations of retrieved data are presented as output. Their format is under system control. It is also possible to request a catalog of the complete content of the data base sorted on any combination of fields. This form of output does not use the search program or the construct.--Harriet R. Meadow.

CONSTRUCTION AND USE OF TABULAR KEYS

I. W. Newell (Dept. of Life Sciences, Univ. of California, Riverside, CA 92502) has just published a paper under the title above, in Pacific Insects, vol. 12, 1970, pp. 25-37. There have been a series of efforts by taxonomic workers to improve upon the standard dichotomous keys traditionally employed by systematists, but they often fall between the cracks. Quite a few of them are almost immediately convertible into computer compatible data matrices, even though the authors do not regard them as steps in the direction. One of the major problems for translation into computer matrices is the diversity of techniques used in coding characters. Newell has designed what he calls "cognitive character codes," designed to provide recognition of the characteristic through key words or letters that act as recognition devices. This produces a variety of symbols in the table, which may accelerate use visually, but lead to difficulties for a computer. A paper by several of us who have been working on this specific problem for some time will appear in Biosciences, probably in February, 1971, and will hopefully provide some basis for development of standard coding methods in the continued development of tabular keys and computer matrices. A shortened abstract of Newell's paper follows.--JAP.

"Tabular keys differ from dichotomous keys in that the statements of characters and variants are separated from the names of the taxa to be identified. This is done by applying cognitive code symbols to the character variants, and recording the distribution of the variants for each of the taxa in a table. The use of Key Groups (KGG) in which carefully selected characters are presented in groups of about five, makes possible the handling of taxa of any size. Tabular keys are open-ended, so expansion of the key can be accomplished by adding new groups of taxa and groups of differentiating characters at the open ends, eliminating the need for complete revision of the key. Selection and coding of characters as well as preparation of the tables and the data card file are described."

AUTOMATED CONSULTING SYSTEM DEVELOPED

An "automated consulting" system for evaluating and interpreting technical farm production and pest control information is being developed cooperatively by the Departments of Agricultural Economics and of Botany and Plant Pathology at Michigan State University. This system offers a practical method for providing computer access for field-based extension personnel. Currently, it is being used on a trial basis by 10 Michigan extension agents who serve commercial agriculture and urban backyard garden enthusiasts.

The use of computers in extension education is considered a major step in establishing a closer link with the local problems of both the rural and city communities. Several computer programs have been designed and tested which provide immediate interpretation, evaluation, and comparison of alternatives avail-

able to a grower in his farming operation. Since the producer supplies information (data) for his own operation, results of the computer analysis are tailored to his individual situation.

Low cost, multi-purpose touch-tone telephone terminals and standard telephone lines are used to gain access to the computer. Data are entered manually into the computer by means of push buttons, or automatically by using punched plastic cards. The automatic feature is particularly useful for feeding in pre-determined digital codes to identify the user and to request the desired computer program.

The novel feature that makes this system feasible and effective is the "spoken" reply received from an Audio Response Unit located near the computer. Once execution of a program begins, the computer will "ask" for the required input values. These are entered using the touch-tone keyboard. After the data have been entered and verified, the proper words are selected from a prerecorded stored vocabulary and returned to the user via telephone. The responding female voice, affectionately known as Synthia by Michigan growers, may be heard over an auxiliary speaker or over the handset's receiver.

Before requesting access to the computer, the user fills in the input portion of an input/output form. Separate forms accompany each program. Qualitative information including names of pesticides, diseases, and symptoms are entered into and retrieved from the computer in coded form. The codes to be used are listed either on the input/output form or on a separate work-sheet. A record of the computer analysis is maintained by writing the results on the output section of the form.

The following example illustrates how the touch-tone telephone is used to identify plant diseases. The computer program was developed by Larry E. Morse, an undergraduate at Michigan State University, for identifying plants. It was converted for operation via touch-tone telephone.

The central feature of this system is a date matrix in which the symptoms exhibited by each disease are stored on disc files in coded form. Currently, the matrices are being developed on a crop basis and include most of the common diseases which occur on each crop in Michigan. Tentative data matrices have been constructed for identifying peach and cherry diseases.

To identify a plant disease specimen, the user must fill out the input/output form for program 09. A supplemental work-sheet listing symptom and disease codes is supplied for each matrix. The user calls the computer located at the University of Michigan and, upon request, enters the input data using the push buttons on the telephone. Almost instantly, the computer responds with the results of the analysis. If identification is not obtained, the analysis may be adjusted based upon symptoms suggested by the computer. If more than one sample is to be identified, the user continues by entering data for the next sample.

Computer diagnosis using the Morse system has the advantage of not being restricted to a fixed sequence as with dichotomous keys. One or more symptoms can be used in the identification process. The computer also lists the most distinguishing symptoms

of the disease. This educational feature helps agents associate symptoms not present in his sample with the disease.

The foregoing is one example of a system which is much more expansive in its capabilities. Currently, 10 programs involving agricultural economics, plant pathology, and soil and crop sciences are operational. In addition to disease identification, the programs include cost analysis, income tax analysis, soil nutrition and herbicide recommendations, sprayer calibration, timing of eradicants for apple scab control, and retrieval of pesticide tolerance and compatibility data.

The system has a potential for a wide number of uses and is relatively inexpensive. Normally, a touch-tone telephone would replace one of the rotary dial type. Rental is $5.00 to $8.00 a month more than the dial telephone. The cost for computer time has been under $2.00 for most of the programs. Long distance toll charges have averaged less than $1.00 per call, free where wide area telephone service (WATS) is available.

Initial support to develop this system was provided by the Kellogg Foundation. Eventually, it is anticipated the user will be billed directly. Experience gained from the Michigan TELFARM and DHIA programs indicate farm businesses will readily pay a fee when a needed service is provided.--A. L. Jones and S. B. Harsh [Reprinted from Phytopathology News, vol. 4, no. 10, Oct., 1970].

RECENT LITERATURE

Adams, R. P. Contour mapping and differential systematics of geographic variation. Systematic Zoology, 19: 385-390.

Crovello, T. J. The effect of change of number of OTU'S in a numerical taxonomic study. Brittonia, 20, 1968, pp. 346-67.

Crovello, T. J. Effects of change of characters and of number of characters in numerical taxonomy. American Midland Naturalist, 81, 1969, pp. 68-86.

Crovello, T. J. Numerical taxonomy: its value to mosquito systematics. Mosquito Systematics Newsletter, 1, 1969, pp. 63-7.

Hale, M. E. & R. Creighton. An automated system for recording exchanges. Flora North America Report no. 32, 1970, 1-9.

Maelzer, D. A. The regression of log N [to base n+1] on log N [to base n] as a test of density dependence: an exercise with computer-constructed density-independent populations. Ecology, 51, pp. 810-822.

Menkhaus, E. J. Time sharing--more glitter than gold. Business Automation, Nov., 1970: 36-42. A peep into the future of time-share computing.

Moss, W. W., D. A. Nickle, and M. G. Emsley. A polynomial treatment of stridulatory file data from several species of the Katydid genus Ceraia .... Notulae Naturae, Academy of Natural Sciences, Philadelphia, no. 432, 1970: 1-13. "Second and fourth degree polynomial equations were fitted to stridulatory file data from several species ...and plotted, using a time-sharing computer terminal." [from authors' abstract.]

Rowell, A. J. The contribution of numerical taxonomy to the genus concept. North Amer. Paleont. Convention, Chicago, 1969 [1970], Proc., Part C, pp. 264-293.

February, 1971
Division of Reptiles and Amphibians
National Museum of Natural History
Washington, D. C. 20560